Lab 5: PCR, Electrophoresis, Expression Vectors, and Cloning!

The polymerase chain reaction (PCR) is a powerful molecular technique that allows scientists to amplify a specific DNA sequence millions of times in just a few hours. The technique was invented by Dr. Kary Mullis in 1983, for which he received the Nobel Prize in Chemistry ten years later. PCR is revolutionizing many areas of genetic research including genetic disease diagnosis, forensic medicine, and molecular evolution.

Within a dividing cell, DNA replication involves a series of enzyme-mediated reactions whose end result is a faithful copy of the entire genome. Enzymes first unwind (denature) the DNA double helix into single strands. Then, an RNA polymerase synthesizes a short stretch of RNA complementary to one of the DNA strands at the start site of replication. This DNA/RNA heteroduplex acts as a “priming site” for the attachment of the DNA polymerase, which then produces the complementary DNA strand.

During PCR, high temperature is used to separate the DNA molecules into single strands, the synthetic sequences of single-stranded DNA (20-30 nucleotides) serve as primers. Two different primer sequences are used to bracket the target region to be amplified. One primer is complementary to one DNA strand at the beginning of the target region, a second primer is complementary to the other strand at the end of the target region.

To perform a PCR reaction, a small quantity of the target DNA is added to a test tube with a buffered solution containing DNA polymerase, short oligonucleotide primers that flank the target region, the four deoxynucleoside triphosphate (dNTPs) building blocks of DNA, and the cofactor Magnesium chloride. The PCR mixture is taken through replication cycles consisting of the following:

1. one to several minutes at 94-96 degrees C, during which the DNA is denatured into single strands.

2. one to several minutes at 50-60 degrees C, during which the primers hybridize or “anneal” (by way of hydrogen bonding) to their complementary sequences on either side of the target sequence. **NOTE: These primers will anneal to the original template DNA AND to any PCR products that were made with them during previous rounds of cycling. This fact allows for the EXPONENTIAL increase in the target sequence with every cycle of PCR!!!**

3. one to several minutes at 72 degrees C, during which the DNA polymerase binds and extends a complementary strand from each primer. The length of this extension time is dependent on the size of the target sequence. By convention, one minute/ 1 kilobase is used.
As amplification proceeds, the DNA sequence between the primers doubles after each cycle. Following thirty such cycles, a theoretical amplification factor of up to 1 billion copies is attained.

Two important innovations were responsible for automating PCR. First, a heat-stable DNA polymerase was isolate from the bacterium *Thermus aquaticus*, which inhabits hot springs, such as the geothermal regions in Yellowstone Park. This enzyme, called “Taq” DNA polymerase, remains active despite repeated heating during many cycles of amplification. Second, DNA thermal cyclers have been invented in which a computer controls the repetitive temperature changes required for PCR. (Information above was taken from http://www.dnalc.org/Shockwave/pcranwhole.html)

Setting up the PCR reaction: Each group should set up a PCR reaction. One member should do the pipeting (to ensure accuracy and precision) but others in the group, make sure every component makes it into the PCR reaction. YOU WILL BE PIPETTING VERY SMALL VOLUMES!

1. Mark a 200 uL PCR tube, containing 22 uL of blue PCR master mix*, with your group’s name. They are tiny tubes!
2. into the PCR tube, carefully combine….
   a. 1 uL of template DNA (Jelly fish cDNA gene sequence for GFP)
   b. 2 ul of primers mix (mix contains both forward and reverse)

   *PCR master mix (contains Taq DNA polymerase, dNTPs, Mg ion, buffering agents)

3. Mix well by pipeting up and down, yet try not to make too many bubbles.
4. Transfer to the thermocycler and amplify using the following settings provided by the instructor. **LET THE INSTRUCTOR KNOW WHEN EVERYONE IS READY!!!** Example settings illustrating the different parts of PCR are shown below. NOTE: THESE MAY BE DIFFERENT FROM THE SETTINGS WE ARE USING…THEY ARE JUST AN EXAMPLE!

<table>
<thead>
<tr>
<th>Setting</th>
<th>Temp degrees C</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>94</td>
<td>5 min</td>
</tr>
<tr>
<td>2.</td>
<td>94</td>
<td>2 min</td>
</tr>
<tr>
<td>3.</td>
<td>42</td>
<td>2 min</td>
</tr>
<tr>
<td>4.</td>
<td>72</td>
<td>2.5 min</td>
</tr>
</tbody>
</table>

repeat steps 2 – 4 through 35 more cycles
5. 72  5 min
6. 4  infinity

Make sure you know what is happening in each step. Discuss it with members of your group.

Procedure for the design of PCR primers.

There is a tremendous amount of genetic and molecular data available for free to access on the internet. One central point is called the National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)). This site contains gene sequences, proteins sequences, useful taxonomy information for nearly all of the genes and proteins so far discovered. You can access scientific articles through this website too.

Log into a web browser and type in the following URL:

To obtain the sequence of green fluorescent protein, go to the imbedded site in the NCBI site. Type the direct link below.


1. Observe the accession number
2. Important information in the heading
3. Gene sequence (CDS) and the translated amino acid sequence.
4. Note: the reported gene sequence is the sequence of deoxyribonucleotides found on the coding strand of the DNA double helix, listed 5' to 3'. What would the complementary strand look like? This would also be the noncoding strand.
5. In PCR, a thermostable DNA polymerase copies the gene sequence and like all DNA polymerases, they can only extend new DNA sequences off of existing 3’OH groups...we need to make a primer (actually two primers) to copy both the coding and noncoding strands in the DNA double helix. (Think of it as DNA replication of a short stretch of DNA.)
6. In the cell, short RNA is used to allow DNA to replicate. In the lab, we design and have produced for us short single stranded DNA primers (oligonucleotides) that will anneal to the appropriate sites on the DNA template.
7. The two primers we generate must bind to the DNA sequence at the beginning and ending regions of the gene of interest.
   a. To do this one primer must be complementary to the 3’ end of the noncoding strand (called the Forward primer)
   b. The other primer must be complementary to the 3’ end of the coding strand (called the Reverse primer)
8. The length of DNA primers is typically around 20-30 nucleotides and have Tm values around 65 degrees C.
For primers greater than 20 nucleotides long, the following formula is used to calculate the Tm.

\[ Tm \,(^\circ C) = 69.3 + [41(\text{sum of the G and C in sequence}) - 650] / \text{primer length} \]

9. Using the gene sequence for GFP...design a set of PCR primers that have a Tm of at least 65\(^\circ\)C.
10. These sequences could be used in a PCR reaction to amplify the gene for GFP from a very dilute template DNA supply, which was originally derived from the Northern Pacific bioluminescent jelly fish *Aequorea Victoria*.

Forward Primer sequence (written 5’ to 3’):

Reverse Primer sequence (written 5’ to 3’)

**Procedure for gel electrophoresis of DNA fragments.** (only 1 gel is needed per class...volunteers?)

1. Prepare 50 mL of a molten 0.8% agarose solution in 1X SB buffer.
   1. The agarose is a solid to be weighed out
   2. The SB buffer comes as a 20X stock.
   3. Combine in a 150 mL beaker and microwave for approximately 40 seconds. DO NOT LET THE SOLUTION BOIL OVER!!
2. Pipet 1 uL of Gelstar nucleic acid gel stain into a 50 mL centrifuge tube. Add the ~50 mL of agarose solution to the tube. Mix by inverting the tube. Caution, this dye is a possible mutagen. WEAR GLOVES WHEN PREPARING OR HANDLING GELS THAT CONTAIN NUCLEIC ACID INTERCALATIN DYES LIKE GELSTAR OR ETHIDIUM BROMIDE!
3. Cast the molten agar into a gel casting try with an 8 well comb in place.
4. Allow the gel to congeal in the fridge to speed up the solidification.
5. Once hardened, move the gel to a gel loading station, add ~500 mL of 1X SB buffer to the gel chamber, covering the gel and wells. Carefully remove the comb at this point by gently pulling straight up. You should have nicely formed wells, ready to load!
Prepare your PCR DNA samples: 5 µL the reaction from each group into a new microcentrifuge tube. Make sure to add loading dye if you need to. Ask the instructor if you are not sure! ALSO KEEP THE REMAINING PORTION OF YOUR PCR REACTION!!!!!! IT IS NEEDED FOR THE NEXT PROCEDURE OF THE LAB, ASSUMING YOUR PCR REACTION WORKED!

1. To each sample, (except for the DNA ladders) add 5 µL of loading dye.
2. Use a micropipet with a fresh tip to add all of the DNA sample/loading dye mixture into the wells of the agarose gel. Expel any air from the tip before loading, and be careful not to push the tip of the pipet through the bottom of the sample well. AND TRY TO MINIMIZE SPILL OVERS!!!!
3. Electrophorese at 275 volts for 12 minutes. You can follow the migration by following the loading dye’s movement in the gel.
4. Observe gel on the UV light box and take picture. Did your PCR work? If not, borrow someone else’s “leftovers”.

**Cloning your PCR product into an Expression Vector**

For those groups that had a successful PCR, CONGRATULATIONS!!! Using the power of the polymerase chain reaction, you have copied the gene for GFP out of the jellyfish genome template DNA and made millions of copies (so much so that you could detect the amplified DNA sequence in only 5 µL of your PCR reaction!) For those groups that weren’t successful…find a nice neighboring group to borrow some of their PCR product for this next section. Your first goal has been accomplished!! You have cloned (copied) the GFP gene. Now we want that gene to be transcribed and translated into GFP protein! To do this, we need to use an **expression vector**. An expression vector is a man-made plasmid DNA that contains several important features. **(KNOW THESE)** They include:

- An origin of replication…Why?______________________________
- A selectable marker…Why? _______________________________
- An inducible promoter site….Why?_________________________
- A transcriptional termination site…Why?____________________
- A site to introduce your gene…Why?_______________________
The circular plasmid expression vector has already been linearized for you. You need to mix the cut vector with your GFP PCR product, then ligate it into the vector, making a new recombinant expression vector.

What kind of DNA enzyme would you use to linearize (cut) a circular piece of DNA like an expression vector? ________________________________________________________________

What kind of DNA enzyme would you use to connect (glue) your PCR product to the vector and “re-close” the vector? ________________________________________________________________

1. Obtain a tube of cut expression vector, DNA ligase, and the rest of your PCR product. Keep them on ice.
2. In a microcentrifuge tube, combine
   a. 10 uL of vector DNA,
   b. 2 uL of PCR-amplified GFP gene, and
   c. 5 uL of ligase enzyme.
   d. Mix by pipetting and incubate at room temperature for **15 minutes**. What is happening in this tube during the 15 minutes?
   What are the possible products of this ligation reaction? There are more than one possible construct! Make sure you know them, discuss it with your partners.
3. While you are incubating your ligation…obtain an agar plate that contains both ampicillin and L-arabinose dissolved in the agar AND an agar plate that contains ampicillin alone. Write your groups names on the bottom of each of the plates and warm them at 37 degrees C for 5 minutes. How are these components going to helps us make GFP protein? ____________________________________________

---

**Quick Transformation and Plating of Competent E. coli**

You have just ligated your PCR gene into the expression vector. Finally, you need to place that recombinant vector into a prokaryotic host to make the protein for us! We are using *E. coli* (a nonpathogenic strain) to make the protein for us. Why do we choose *E. coli*? 1. They divide rapidly and grow quickly 2. they are capable of taking up foreign DNA like expression vectors (a process called transformation). 3. They can be induced to make lots of proteins.

In order to carry out a transformation experiment, you need to make your *E. coli* “competent”. That is, to make them able to take up foreign DNA. This is done in the lab by treating the *E. coli* with a divalent cation like Ca\(^{2+}\) ions. This has already been done for you. You need to obtain the treated and LIVE *E. coli*, place them in a tube and add your ligated expression vector. We do a quick heatshock to allow the bacteria to take up the vector, then plate the transformed bacteria on an agar plate to grow overnight.

**IMPORTANT**: YOU WILL BE USING LIVE BACTERIA...WEAR GLOVES, STERILIZE YOUR WORK AREA BEFORE AND AFTER WITH ALCOHOL, DISPOSE OF WASTE IN THE PROPER SPOT!!!!!

1. Clean your area with alcohol
2. put on gloves
3. Obtain a tube of competent bacteria already in a transformation tube
4. place immediately on ice
5. Add all of your ligation reaction (~10 uL) to the bacteria and mix gently by stirring with a pipette tip on your micropipettor.
6. incubate on ice for 5 minutes
7. Heatshock the tube at 42 degrees C for **30 SECONDS**, RETURN TO ICE FOR 5 MINUTES
8. Pipette ~50 uL of the transformed bacteria onto each agar plate.
9. aseptically spread the bacteria on the surface of the agar and allow to air dry
10. Sterilize your work area with alcohol.
11. turn upside down and incubate overnight in incubator
12. The next day, remove the plates from the incubator and place them in the fridge.
13. We’ll examine your results next time.

**The Arabinose Operon: Control of our Expression Vector**

The Arabinose Operon

The *ara* operon codes for three enzymes that are required to catalyze the metabolism of arabinose. They are under the control of the arabinose promoter system, also called pBAD. When these genes need to be turned on, transcription factors bind the pBAD and recruit RNA polymerase to begin transcription. The genes controlled by pBAD include:

- **Arabinose isomerase** - encoded by *araA* - converts arabinose to ribulose
- **Ribulokinase** - encoded by *araB* -- phosphorylates ribulose
- **Ribulose-5-phosphate epimerase** - encoded by *araD* -- converts ribulose-5-phosphate to xylulose-5-phosphate which can then be metabolized via the pentose phosphate pathway.

The three structural genes are arranged in an operon that is regulated by the *araC* gene product. The araC protein is a main "gatekeeper" protein for this promoter system. It helps to logically control the expression of *araA*, *araB*, and *araD* genes. There are four important regulatory sites as shown in the following diagram:
• *araO₁* is an operator site. *AraC* binds to this site and represses its own transcription from the \( P_C \) promoter. In the presence of arabinose, however, *AraC* bound at this site helps to activate expression of the \( P_{BAD} \) promoter.

• *araO₂* is also an operator site. *AraC* bound at this site can simultaneously bind to the *araI* site to repress transcription from the \( P_{BAD} \) promoter.

• *araI* is also the inducer site. *AraC* bound at this site can simultaneously bind to the *araO₂* site to repress transcription from the \( P_{BAD} \) promoter. In the presence of arabinose, however, *AraC* bound at this site helps to activate expression of the \( P_{BAD} \) promoter.

• *CRP* binds to the *CRP* binding site. It does not directly assist RNA polymerase to bind to the promoter in this case. Instead, in the presence of arabinose, it promotes the rearrangement of *AraC* when arabinose is present from a state in which it represses transcription of the \( P_{BAD} \) promoter to one in which it activates transcription of the \( P_{BAD} \) promoter.

Regulation of the arabinose operon is both **positively** and **negatively** controlled.

When arabinose is absent, there is no need to express the structural genes. *AraC* does this by binding simultaneously to *araI* and *araO₂*. (the *AraC* protein forms a protein dimer-two sticks together-to “cross link the DNA loop upstream of the pBAD genes). It is shown as the two circular objects in the diagram below. As a result the intervening DNA is **looped**. An abundance of glucose (the preferred sugar in most organisms) helps to silence this operon by keeping the CRP region “open”. This should make sense to you. If bacteria are provided with glucose, then backup operons such as this one should be silent. This is negative regulation. These events block access to the \( P_{BAD} \) promoter which is, in any case, a very weak promoter (unlike the *lac* promoter):
**AraC** also prevents its own expression. Thus, it is an autoregulator of its own expression. This makes sense; there is no need to over-express **AraC**. If the concentration falls too low then transcription of **araC** resumes until the amount of **AraC** is sufficient to prevent more transcription again.

When arabinose is present, it binds to **AraC** and allosterically induces it to bind to **araI** instead **araO**. If **glucose** is also **absent**, then the presence of **CRP** bound to its site between **araO** and **araI** helps to break the DNA loop and also helps **AraC** to bind to **araI**:

![Diagram of the ara operon](image)

The **ara** operon demonstrates both negative and positive control. It shows a different function for **CRP**. It also shows how a protein can act as a switch with its activity being radically altered upon the binding of a small molecule.
Questions

1. What is a reason for bioluminescence in the ocean? Briefly, describe an example.

2. Sketch a picture of your recombinant expression vector containing and labeling the important regions needed for this construct to perform as expected.

3. Explain why GFP is turned on in the presence of L-arabinose.

4. What would be the explanation for colonies that grew on our plates, but were not fluorescent green? Give a probable reason.

5. If a mutation occurred in the araC gene that produced an araC protein that could NOT bind to the (operator 2 ) O_2 region, what would be the phenotype of the *E. coli*? Explain.

6. Observe the glo-fish in the tank in LS-224. Compare their fluorescence to the wild type zebra danio in the tank.
   a. What parts of the fish are glowing?
   Visit the following website, and read a press release for Glo-fish. Type in [http://www.glofish.com/pressreleases/gf.pr.20031121.pdf](http://www.glofish.com/pressreleases/gf.pr.20031121.pdf) Write a short paragraph sharing your thoughts and opinions about transgenic organisms. Address the following: Should transgenic organisms be created? Should they be sold as pets? When would it be appropriate to create a transgenic organism, if at all?

Make sure you understand how the pBAD expression vector works!